Background

The City of Seattle Department of Information Technology (DoIT) contracted with a consultant team (Applied Inference, Pacific Market Research and Andrew Gordon of the UW) to assess the current level of information technology access and literacy among Seattle's residents, to explore residents' perceptions about information technology and to assist in the cable franchise renewal process and will be used as part of the community needs assessment. City staff from DoIT and the Citizens Telecommunications and Technology Advisory Board (CTTAB) developed a set of indicators to measure a broad range of impacts that information technology is having on the Seattle area. This survey is intended to update and, when possible, to compare with a similar survey conducted in 2000.

Methods

DoIT and CTTAB staff with the advice of their cable franchise consultants collaborated with the survey consulting team to develop and refine the current 19-minute survey, to be administered by telephone to 1000 random Seattle residents.

Survey Sampling, Inc. provided a random sample of Seattle area telephone numbers. As the interviewing process progressed, it became evident that ethnic minority respondents were under represented in the sample of completed interviews. In response, a targeted sample of telephone numbers was ordered focusing on Seattle ZIP codes with a higher incidence of ethnic minority households.

Overall, 4894 telephone calls were made that resulted in contact with an individual. Of these 1000 resulted in completed surveys and 3894 were terminated early. Table 1 summarizes the reasons for termination.

Table 1. Disposition of terminated calls

	Number	Percent of calls
Completes	**1000**	20%
Unqualified incompletes	**425**	9%
Not 18 or older	206	
Not a Seattle resident	154	
No such person	14	
Claims previous interview	20	
Over quota (ethnicity, gender)	31	
Refusals, before screener	**2594**	53%
Never call this number	33	
Screener refusal/break off	35	
Hard refusals	1048	
Soft refusals	1478	
Callback (scheduled or not)	**440**	9%
Qualified refusal	**39**	1%
Communication/language	**396**	8%
Spanish	104	
Asian	112	
Other/unknown	135	
Other communication	45	
Total	4894	100%

Table 1 shows that about half of those contacted did not wish to participate in any survey. It is not known how many of these individuals might have been qualified. Nine percent of the calls were terminated because the person was unqualified[1] for the survey, another eight percent because of a language or communication barrier and for another nine percent, the respondent asked to be re-contacted at a better time. Very few (39) refused to continue after qualifying. Of all the calls reaching a household, 20% completed the survey. Of those known to be qualified, 68% completed the survey.

In addition to those calls detailed in Table 1, other phone numbers were dialed. Table 2 summarizes the disposition of those calls.

Table 2. Disposition of other phone numbers dialed

	Number	Percent
No answer/ busy/ answering machine	2000	21%
Non working numbers	5554	59%
Non-residential numbers	1143	12%
Other phone problems (fax/modem)	633	7%
Total	9330	100%

Tables 2 shows that of the calls where the interviewer did not reach a member of a household, 59% were non working numbers. Another 21% of the numbers rang, but was not answered. Twelve percent of the numbers were non residential and seven percent were fax or modem numbers.

Weights

Because of the targeted sampling discussed above the geographic distribution of survey respondents was dissimilar to the distribution of the city's residents. To correct for this, weights were calculated so that individuals from undersampled ZIP codes would be counted more heavily and those from oversampled ZIP codes would be counted less heavily. In subsequent calculations, respondents were further balanced according to ethnicity, age and income. Tables 3 and 4 display the distribution of the sample at each stage of the weighting procedure, compared with the 2000 U.S. Census data.

[1] Disqualified individuals would be those younger than 18, or living outside the city of Seattle.

Table 3. Distribution of respondents by Seattle ZIP code after iterative adjustments in weights

			Adjusted by			
	Seattle population[2]	Unweighted sample	ZIP code	ZIP code and ethnicity	ZIP code, ethnicity and age	ZIP code, ethnicity, age and income
ZIP code						
98101	1.6%	1.5%	1.6%	1.5%	1.4%	1.4%
98102	3.5%	2.6%	3.5%	3.4%	3.4%	3.4%
98103	7.0%	7.1%	7.0%	6.6%	6.6%	6.5%
98104	2.3%	1.1%	2.3%	2.5%	2.2%	2.3%
98105	6.7%	5.1%	6.7%	6.6%	7.5%	7.3%
98106	3.2%	3.3%	3.2%	3.5%	3.5%	3.5%
98107	3.5%	3.3%	3.5%	3.4%	3.7%	3.7%
98108	3.0%	3.7%	3.0%	3.6%	3.6%	3.5%
98109	2.8%	2.1%	2.8%	2.5%	2.4%	2.4%
98112	3.3%	2.1%	3.3%	3.1%	3.2%	3.2%
98115	6.8%	7.3%	6.8%	6.1%	5.9%	5.9%
98116	3.4%	2.8%	3.4%	3.2%	3.0%	3.0%
98117	4.7%	4.4%	4.7%	4.3%	4.2%	4.2%
98118	5.7%	17.5%	5.7%	7.1%	7.1%	7.0%
98119	3.3%	3.7%	3.3%	3.0%	3.0%	2.9%
98121	1.6%	1.8%	1.6%	1.5%	1.5%	1.4%
98122	4.7%	4.1%	4.7%	5.0%	5.0%	5.2%
98125	5.5%	5.4%	5.5%	5.2%	5.2%	5.2%
98126	3.7%	3.0%	3.7%	3.4%	3.3%	3.4%
98133	6.7%	3.0%	6.7%	7.0%	7.8%	7.8%
98134	0.1%	0.3%	0.1%	0.1%	0.1%	0.1%
98136	2.6%	1.6%	2.6%	2.5%	2.4%	2.3%
98144	4.7%	6.1%	4.7%	5.3%	5.1%	5.3%
98177	3.6%	1.3%	3.6%	3.1%	2.7%	2.7%
98178	3.2%	3.9%	3.2%	3.9%	3.8%	3.9%
98199	3.0%	1.9%	3.0%	2.6%	2.5%	2.5%

[2] Seattle population values based on 2000 U.S. Census figures for Seattle

Table 4. Distribution of respondents by demographic categories after iterative adjustments in weights

			Adjusted by			
	Seattle population	Unweighted sample	ZIP code	ZIP code and ethnicity	ZIP code, ethnicity and age	ZIP code, ethnicity, age and income
Gender						
Male	49.8%	50%	51.5%	50.5%	50.4%	50.4%
Female	50.2%	50%	48.5%	49.5%	49.6%	49.6%
Race/ethnicity						
African American	8.4%	9%	5.6%	8.4%	8.1%	8.1%
Asian / Pacific Islander	13.5%	9%	7.9%	13.5%	15.0%	15.4%
Caucasian	69.5%	76%	80.7%	69.5%	67.7%	67.2%
Hispanic / Latino	5.2%	3%	2.2%	5.2%	5.7%	5.8%
Native American / American Indian	1.0%	1%	1.7%	1.0%	1.0%	1.0%
Other	2.4%	1%	1.3%	2.4%	2.4%	2.4%
Age category						
18-24	14.3%	8%	8.4%	11.0%	14.3%	14.5%
25-34	26.3%	22%	22.3%	24.8%	26.3%	26.0%
35-49	30.3%	31%	29.3%	32.0%	30.3%	30.5%
50-64	16.7%	26%	25.8%	26.8%	16.7%	16.4%
65+	12.4%	13%	13.2%	13.7%	12.4%	12.5%
Income category[3]						
Extremely low (>30% of median)	12.9%	13%	11.9%	13.4%	13.8%	12.9%
Low (30% - 49%)	14.6%	12%	11.5%	12.2%	12.8%	14.6%
Moderate (50% - 79%)	18.4%	14%	13.7%	15.2%	15.8%	18.4%
Middle (80% - 94%)	13.2%	18%	16.8%	16.9%	16.9%	13.2%
Upper middle (95% - 119%)	15.8%	15%	14.1%	14.9%	14.9%	15.8%
Upper/ high upper (120%+)	25.2%	27%	27.8%	27.4%	25.9%	25.2%

Inferential analyses, usually factorial analysis of variance or two-way frequency distributions with a chi-square statistic, are conducted where appropriate assumptions are met. These analyses were computed without weights; however, weighted percentages and means were reported. All differences reported are statistically significant at $p<.05$, unless otherwise noted in the narrative.

[3] Population percentages based on regional figures for income category

Limitations

Telephone surveys have fundamental limitations:

- The findings represent only those households that have a working telephone. According to the 2000 U.S. Census, this number is high in Seattle (98.9% of Seattleites have working telephones at home), so this is not likely to present a substantial bias.
- When conducted in English, telephone surveys require that a qualified person (in this case, someone 18 or older) be able to speak English well enough to participate. Also according to the 2000 U.S. Census, 9.3% of Seattleites speak English less than "very well." This could be more problematic and might lead to under-representation of important groups in this survey, and an overestimation of the indicators.
- People who agree to participate in a telephone survey and who persevere through it may be different in other ways from people who refuse to participate at all or who do not complete.

These are some of the ways in which the sample might be biased. That is, the sample may be made up of individuals who may not be representative of all the community's residents. The practice of applying weights to certain subgroups is an effort to balance the sample to make it more similar in certain characteristics to the population, but it cannot make up for subgroups that are missing entirely.

A separate concern is the accuracy and representative-ness of the responses themselves. This issue is addressed with the concept of the confidence interval. This concept is based on the idea that any sample is unlikely to provide responses that are the exact true population values. As the sample size grows, the sample responses become increasingly likely to be closer to the population values. In a survey of 1000 adults, statements about the population are made with 95% confidence that the values reported are within three percentage points of the true population values (± 3%). Figure 1 below shows that 83% of the respondents report having a home computer. Putting this into the context of a confidence interval, since this is based on the sample of 1000, we can be 95% sure that between 80% and 86% of Seattle's residents have home computer access. When conclusions are being drawn about subgroups in the population, the confidence interval grows, so that percentages representing a subgroup of 100 would have a confidence interval of ± 10%. (For inferential statistics, when a significant difference is found between subgroups, we are at least 95% certain that the difference found in the sample is representative of a similar difference in the population and not due to chance fluctuations in the data.)

Combining this issue (non-representative-ness of responses) with the issue of bias, perhaps corrected by applying heavier weights to certain subgroups, can have the effect of exaggerating a non-representative sample in a way that could be difficult to detect.

When interpreting these findings, it is important to keep these limitations in mind and look for patterns in the findings, remembering both that some voices are likely to be missing from this report and those that are present might not accurately represent others in their subgroup. On the other hand, this is a large sample and the findings are consistent, fit a pattern and seem to make sense.